Champ: Controllable and Consistent Human Image Animation with 3D Parametric Guidance

Year: 2024 | Citations: 211 | Authors: Shenhao Zhu, Junming Leo Chen, Zuozhuo Dai, Yinghui Xu, Xun Cao

---

Abstract

In this study, we introduce a methodology for human image animation by leveraging a 3D human parametric model within a latent diffusion framework to enhance shape alignment and motion guidance in curernt human generative techniques. The methodology utilizes the SMPL(Skinned Multi-Person Linear) model as the 3D human parametric model to establish a unified representation of body shape and pose. This facilitates the accurate capture of intricate human geometry and motion characteristics from source videos. Specifically, we incorporate rendered depth images, normal maps, and semantic maps obtained from SMPL sequences, alongside skeleton-based motion guidance, to enrich the conditions to the latent diffusion model with comprehensive 3D shape and detailed pose attributes. A multi-layer motion fusion module, integrating self-attention mechanisms, is employed to fuse the shape and motion latent representations in the spatial domain. By representing the 3D human parametric model as the motion guidance, we can perform parametric shape alignment of the human body between the reference image and the source video motion. Experimental evaluations conducted on benchmark datasets demonstrate the methodology's superior ability to generate high-quality human animations that accurately capture both pose and shape variations. Furthermore, our approach also exhibits superior generalization capabilities on the proposed in-the-wild dataset. Project page: https://fudan-generative-vision.github.io/champ.